

CITED BY APPLICANT

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
27. Dezember 2002 (27.12.2002)

PCT

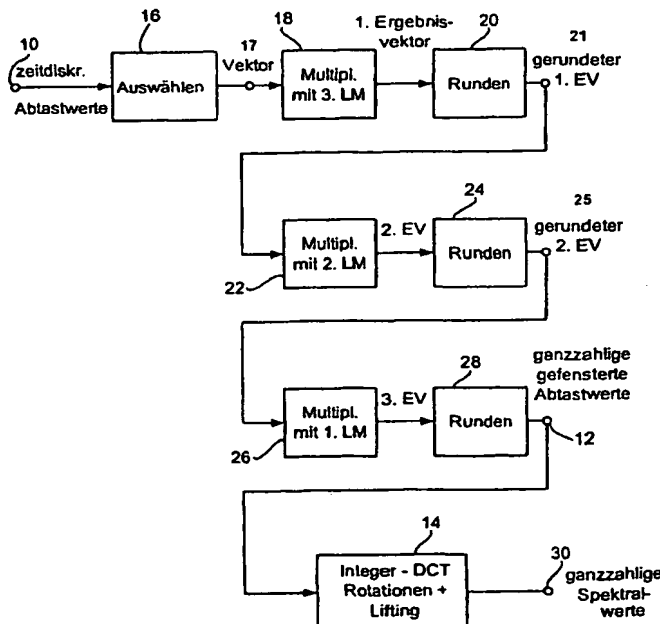
(10) Internationale Veröffentlichungsnummer
WO 02/103684 A1

- (51) Internationale Patentklassifikation: **G10L 19/02, G06F 17/14**
- (21) Internationales Aktenzeichen: **PCT/EP02/05865**
- (22) Internationales Anmeldedatum: **28. Mai 2002 (28.05.2002)**
- (25) Einreichungssprache: **Deutsch**
- (26) Veröffentlichungssprache: **Deutsch**
- (30) Angaben zur Priorität: **101 29 240.6 18. Juni 2001 (18.06.2001) DE**
- (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): **FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUNG E. V. (DE/DE); Leonrodstrasse 54, 80636 München (DE).**
- (72) Erfinder; und
- (75) Erfinder/Anmelder (nur für US): **GEIGER, Ralf (DE/DE); Weide 7, 98704 Wümbach (DE). SPORER, Thomas (DE/DE); Wilhelmshavener Strasse 15, 90766 Fürth (DE). KOLLER, Jürgen (DE/DE); Weide 7, 98704 Wümbach (DE). BRANDENBURG, Karlheinz (DE/DE); Haagstrasse 32, 91054 Erlangen (DE). HERRE, Jürgen (DE/DE); Am Eichengarten 11, 91054 Buckenhof (DE).**

[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD AND DEVICE FOR PROCESSING TIME-DISCRETE AUDIO SAMPLED VALUES

(54) Bezeichnung: VERFAHREN UND VORRICHTUNG ZUM VERARBEITEN VON ZEITDISKRETEN AUDIO-ABTASTWERTEN



- | | |
|------------------------------------|---------------------------------------|
| 10... TIME-DISCRETE SAMPLED VALUES | 24... ROUNDING |
| 16... SELECTING | 25... ROUNDED 2.EV |
| 17... VECTOR | 26... MULTIPLY BY 1.LM |
| 18... MULTIPLY BY 3. LM | 28... ROUNDING |
| 1... RESULT VECTOR | 12... INTEGER WINDOWED SAMPLED VALUES |
| 20... ROUNDING | 14... INTEGER-DCT ROTATIONS+LIFTING |
| 21... ROUNDED 1.EV | 30... INTEGER SPECTRAL VALUES |
| 22... MULTIPLY BY 2.LM | |

(57) Abstract: In order to obtain an integer transformation, which provides integer initial values, the TDAC (time domain aliasing cancellation) function of an MCDT (modified discrete cosine transform) is executed explicitly in the time domain before the forward transformation. In the event of overlapping windows, this leads to a Givens rotation that can be represented by lifting matrices. Time-discrete sampled values of an audio signal are initially combined in pairs to form a vector in order to be sequentially subjected to the action of a lifting matrix after the other. According to the invention, a rounding step (20, 24, 28) is executed by using a lifting matrix after each multiplication (18, 22, 26) of a vector whereby resulting only in integers on the output side. A spectral representation with integer spectral values can be obtained by transforming (14) the windowed integer sampled values with an integer transformation. The inverse imaging with an inverse rotation matrix and with corresponding inverse lifting matrices leads to a precise reconstruction. The inventive concept provides a loss-free transformation, which can be directly coupled to an entropy coder without quantization in order to obtain a windowing method and transformation method that can be implemented advantageously with regard to hardware.

[Fortsetzung auf der nächsten Seite]

WO 02/103684 A1

BEST AVAILABLE COPY



(74) **Anwälte:** SCHOPPE, Fritz usw.; Schoppe, Zimmermann, Stöckeler & Zinkler, Postfach 71 08 67, 81458 München (DE).

Veröffentlicht:

-- mit internationalem Recherchenbericht

(81) **Bestimmungsstaaten (national):** JP, US.

(84) **Bestimmungsstaaten (regional):** europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(57) Zusammenfassung: Um eine Ganzzahl-Transformation zu erhalten, die ganzzahlige Ausgangswerte liefert, wird die TDAC-Funktion einer MDCT explizit im Zeitbereich vor der Hin-Transformation ausgeführt. Bei überlappenden Fenstern führt dies zu einer Givens-Rotation, die durch Lifting-Matrizen dargestellt werden kann, wobei zeitdiskrete Abtastwerte eines Audiosignals zunächst paarweise zu einem Vektor zusammengefasst werden, um sequentiell mit einer Lifting-Matrix nach der anderen beaufschlagt zu werden. Erfindungsgemäss wird nach jeder Multiplikation (18, 22, 26) eines Vektors mit einer Lifting-Matrix ein Rundungsschritt (20, 24, 28) durchgeführt, so dass sich ausgangsseitig lediglich Ganzzahlen ergeben. Durch Transformieren (14) der gefensterten ganzzahligen Abtastwerte mit einer ganzzahligen Transformation kann eine spektrale Darstellung mit ganzzahligen Spektralwerten erreicht werden. Die inverse Abbildung mit einer inversen Drehmatrix und entsprechenden inversen Lifting-Matrizen führt zu einer genauen Rekonstruktion. Das erfindungsgemässe Konzept liefert eine verlustlose Transformation, die ohne Quantisierung unmittelbar mit einem Entropie-Codierer gekoppelt werden kann, um ein Fensterungs- und Transformations-Verfahren zu erhalten, das hardwaremässig günstig implementiert werden kann.

Verfahren und Vorrichtung zum Verarbeiten von zeitdiskreten Audio-Abtastwerten

5

Beschreibung

Die vorliegende Erfindung bezieht sich auf die Audiocodierung und insbesondere auf Verfahren und Vorrichtungen zum Verarbeiten von zeitdiskreten Audio-Abtastwerten, um ganzzahlige Ausgangswerte zu erhalten.

Moderne Audiocodierverfahren, wie z. B. MPEG Layer3 (MP3) oder MPEG AAC verwenden Transformationen wie beispielsweise die sogenannte modifizierte diskrete Cosinustransformation (MDCT), um eine blockweise Frequenzdarstellung eines Audiosignals zu erhalten. Ein solcher Audiocodierer erhält üblicherweise einen Strom von zeitdiskreten Audio-Abtastwerten. Der Strom von Audio-Abtastwerten wird gefenstert, um einen gefensterten Block von beispielsweise 1024 oder 2048 gefensterten Audio-Abtastwerten zu erhalten. Zur Fensterung werden verschiedene Fensterfunktionen eingesetzt, wie z. B. ein Sinus-Fenster, etc.

Die gefensterten zeitdiskreten Audio-Abtastwerte werden dann mittels einer Filterbank in eine spektrale Darstellung umgesetzt. Prinzipiell kann hierzu eine Fourier-Transformation, oder aus speziellen Gründen eine Abart der Fourier-Transformation, wie z. B. eine FFT oder, wie es ausgeführt worden ist, eine MDCT eingesetzt werden. Der Block von Audio-Spektralwerten am Ausgang der Filterbank kann dann je nach Bedarf weiter verarbeitet werden. Bei den oben bezeichneten Audio-Codierern folgt eine Quantisierung der Audio-Spektralwerte, wobei die Quantisierungsstufen typischerweise so gewählt werden, daß das durch das Quantisieren eingeführt Quantisierungsrauschen unterhalb der psychoakustischen Maskierungsschwelle liegt, d. h. "wegmaskiert" wird. Die Quantisierung ist eine verlustbehaftete Codierung. Um eine weitere Datenmengenreduktion zu erhal-

ten, werden die quantisierten Spektralwerte anschließend beispielsweise mittels einer Huffman-Codierung Entropie-codiert. Durch Hinzufügen von Seiteninformationen, wie z. B. Skalenfaktoren etc. wird aus den Entropie-codierten
5 quantisierten Spektralwerten mittels eines Bitstrom-Multiplexers ein Bitstrom gebildet, der gespeichert oder übertragen werden kann.

Im Audio-Decodierer wird der Bitstrom mittels eines
10 Bitstrom-Demultiplexers in codierte quantisierte Spektralwerte und Seiteninformationen aufgeteilt. Die Entropie-codierten quantisierten Spektralwerte werden zunächst Entropie-decodiert, um die quantisierten Spektralwerte zu erhalten. Die quantisierten Spektralwerte werden dann in-
15 vers quantisiert, um decodierte Spektralwerte zu erhalten, die Quantisierungsrauschen aufweisen, das jedoch unterhalb der psychoakustischen Maskierungsschwelle liegt und daher unhörbar sein wird. Diese Spektralwerte werden dann mittels eines Synthese-Filterbank in eine zeitliche Darstellung umgesetzt, um zeitdiskrete decodierte Audio-Abtastwerte zu
20 erhalten. In der Synthese-Filterbank muß ein zum Transformations-Algorithmus inverser Transformations-Algorithmus eingesetzt werden. Außerdem muß nach der Frequenz-Zeit-Rücktransformation das Fenstern rückgängig gemacht werden.

25 Um eine gute Frequenzselektivität zu erreichen, verwenden moderne Audio-Codierer typischerweise eine Block-Überlappung. Ein solcher Fall ist in Fig. 4a dargestellt. Zunächst werden beispielsweise 2048 zeitdiskrete Audio-
30 Abtastwerte genommen und mittels einer Einrichtung 402 gefenstert. Das Fenster, das die Einrichtung 402 verkörpert, hat eine Fensterlänge von $2N$ Abtastwerten und liefert ausgangsseitig einen Block von $2N$ gefensterten Abtastwerten. Um eine Fensterüberlappung zu erreichen, wird mittels einer
35 Einrichtung 404, die lediglich aus Übersichtlichkeitsgründen in Fig. 4a getrennt von der Einrichtung 402 dargestellt ist, ein zweiter Block von $2N$ gefensterten Abtastwerten gebildet. Die in die Einrichtung 404 eingespeisten 2048 Ab-

tastwerte sind jedoch nicht die an das erste Fenster unmittelbar anschließenden zeitdiskreten Audio-Abtastwerte, sondern beinhalten die zweite Hälfte der durch die Einrichtung 402 gefensterten Abtastwerte und beinhalten zusätzlich lediglich 1024 "neue" Abtastwerte. Die Überlappung ist durch eine Einrichtung 406 in Fig. 4a symbolisch dargestellt, die einen Überlappingsgrad von 50% bewirkt. Sowohl die durch die Einrichtung 402 ausgegebenen $2N$ gefensterten Abtastwerte als auch die durch die Einrichtung 404 ausgegebenen $2N$ gefensterten Abtastwerte werden dann mittels einer Einrichtung 408 bzw. 410 dem MDCT-Algorithmus unterzogen. Die Einrichtung 408 liefert gemäß dem bekannten MDCT-Algorithmus N Spektralwerte für das erste Fenster, während die Einrichtung 410 ebenfalls N Spektralwerte liefert, jedoch für das zweite Fenster, wobei zwischen dem ersten Fenster und dem zweiten Fenster eine Überlappung von 50% besteht.

Im Decodierer werden die N Spektralwerte des ersten Fensters, wie es in Fig. 4b gezeigt ist, einer Einrichtung 412, die eine inverse modifizierte diskrete Cosinustransformation durchführt, zugeführt. Dasselbe gilt für die N Spektralwerte des zweiten Fensters. Diese werden einer Einrichtung 414 zugeführt, die ebenfalls eine inverse modifizierte diskrete Cosinustransformation durchführt. Sowohl die Einrichtung 412 als auch die Einrichtung 414 liefern jeweils $2N$ Abtastwerte für das erste Fenster bzw. $2N$ Abtastwerte für das zweite Fenster.

In einer Einrichtung 416, die in Fig. 4b mit TDAC (TDAC = Time Domain Aliasing Cancellation) bezeichnet ist, wird die Tatsache berücksichtigt, daß die beiden Fenster überlappend sind. Insbesondere wird ein Abtastwert y_1 der zweiten Hälfte des ersten Fensters, also mit einem Index $N+k$, mit einem Abtastwert y_2 aus der ersten Hälfte des zweiten Fensters, also mit einem Index k summiert, so daß sich ausgangsseitig, also im Decodierer, N decodierte zeitliche Abtastwerte ergeben.

Es sei darauf hingewiesen, daß durch die Funktion der Einrichtung 416, die auch als Add-Funktion bezeichnet wird, die in dem durch Fig. 4a schematisch dargestellten Codierer durchgeführte Fensterung gewissermaßen automatisch berücksichtigt wird, so daß in dem durch Fig. 4b dargestellten Decodierer keine explizite "inverse Fensterung" stattzufinden hat.

Wenn die durch die Einrichtung 402 oder 404 implementierte Fensterfunktion mit $w(k)$ bezeichnet wird, wobei der Index k den Zeitindex darstellt, bezeichnet wird, so muß die Bedingung erfüllt sein, daß das Fenstergewicht $w(k)$ im Quadrat addiert zu dem Fenstergewicht $w(N+k)$ im Quadrat zusammen 1 ergibt, wobei k von 0 bis $N-1$ läuft. Wenn ein Sinus-Fenster verwendet wird, dessen Fenster-Gewichtungen der ersten Halbwelle der Sinus-Funktion folgen, so ist diese Bedingung immer erfüllt, da das Quadrat des Sinus und das Quadrat des Cosinus für jeden Winkel zusammen den Wert 1 ergeben.

Nachteilig an dem in Fig. 4a beschriebenen Fenster-Verfahren mit anschließender MDCT-Funktion ist die Tatsache, daß die Fensterung durch Multiplikation eines zeitdiskreten Abtastwerts, wenn an ein Sinus-Fenster gedacht wird, mit einer Gleitkommazahl erreicht wird, da der Sinus eines Winkels zwischen 0 und 180 Grad abgesehen von dem Winkel 90 Grad keine Ganzzahl ergibt. Auch wenn ganzzahlige zeitdiskrete Abtastwerte gefenstert werden, entstehen nach dem Fenstern also Gleitkommazahlen.

Daher ist, auch wenn kein psychoakustischer Codierer verwendet wird, d. h. wenn eine verlustlose Codierung erreicht werden soll, am Ausgang der Einrichtungen 408 bzw. 410 eine Quantisierung notwendig, um eine einigermaßen überschaubare Entropie-Codierung durchführen zu können.

Wenn also bekannte Transformationen, wie sie anhand von Fig. 4a betrieben worden sind, für ein verlustloses Audio-codieren eingesetzt werden soll, muß entweder eine sehr

feine Quantisierung eingesetzt werden, um den resultierenden Fehler aufgrund der Rundung der Gleitkommazahlen vernachlässigen zu können, oder das Fehlersignal muß zusätzlich beispielsweise im Zeitbereich codiert werden.

5

Ferner haben digitale Signalprozessoren üblicherweise einen Akkumulator mit größerer Wortlänge als die übliche Operatorlänge, um zu viele Rundungsoperationen zu vermeiden. Die Verwendung von schnellen Algorithmen für die Implementation einer Filterbank führt typischerweise zu dem Bedarf, Zwischenergebnisse zu speichern, die in späteren Schritten verwendet werden sollen. Die Zwischenergebnisse müssen auf die Operatorgenauigkeit gerundet werden oder müssen in zwei Speicheroperationen aufgeteilt werden. Typischerweise akkumulieren sich die Rundungsfehler mehrerer Verarbeitungsschritte. Wenn ferner berücksichtigt wird, daß die meisten Gleitkomma-DSPs eine Wortlänge von 32 Bit mit einer Mantisse von nur 24 Bit haben, ist es offensichtlich, was mit Eingangssignalen mit einer Genauigkeit von 24 Bit passiert.

20

Sowohl eine zu feine Quantisierung als auch das zusätzliche Codieren des Fehlersignals als Alternative resultieren in einem Codierer mit erhöhtem Rechenaufwand und erhöhter Komplexität, und damit auch in einem entsprechend aufwendigeren Decodierer. Insbesondere der Decodierer, der, wenn an die Verbreitung von Musik beispielsweise über das Internet gedacht wird, ein Massenprodukt ist, muß aufgrund dieser Tatsache preisgünstig sein, um sich gegenüber anderen Codierern am Markt durchsetzen zu können. Mit dieser Anforderung ist es oftmals nicht vereinbar, eine sehr feine Quantisierung zu verwenden oder ein zusätzliches Fehlersignal zu codieren, da der zusätzliche Mehraufwand zu höheren Decodiererkosten führt.

30

Gleichzeitig ist es auf Codiererseite auf dem konkurrenzintensiven Markt für Audiocodierer oftmals nicht tolerabel, zu große Datenmengen zu erzeugen. In anderen Worten ausgedrückt ist es von wesentlicher Bedeutung, einen möglichst

35

hohen Kompressionsfaktor zu erreichen, da oftmals Bandbreite-begrenzte Netze existieren, welche dazu führen, daß ein zu gering komprimiertes Audiostück eine zu große Übertragungsdauer über ein solches Netz hat, was beim Kunden unmittelbar dazu führt, daß er sich für ein anderes Produkt mit höherer Datenkompression und damit geringerer Übertragungszeit entscheidet.

Die Aufgabe der vorliegenden Erfindung besteht darin, ein Codierer/Decodiererkonzept zu schaffen, das für die verlustlose Codierung geeignet ist und gleichzeitig eine hohe Datenkomprimierung bei vertretbarem Aufwand liefert.

Diese Aufgabe wird durch ein Verfahren zum Verarbeiten von zeitdiskreten Audio-Abtastwerten nach Patentanspruch 1, durch ein Verfahren zum Invers-Verarbeiten von ganzzahligen Werten nach Patentanspruch 17, durch eine Vorrichtung zum Verarbeiten von zeitdiskreten Audio-Abtastwerten nach Patentanspruch 22 oder durch eine Vorrichtung zum Invers-Verarbeiten von ganzzahligen Werten nach Patentanspruch 23 gelöst.

Der vorliegenden Erfindung liegt die Erkenntnis zugrunde, daß das Entstehen von Gleitkommawerten im Schritt des Fensters dadurch vermieden werden kann, wenn die TDAC-Operation explizit im Zeitbereich, d. h. vor dem Ausführen einer Transformation, durchgeführt wird. Dies wird dadurch erreicht, daß im Gegensatz zum Stand der Technik bereits vor der Transformation die Überlappung berücksichtigt wird, und zwei zeitdiskrete Abtastwerte aus unterschiedlichen Vierteln eines Fensters verarbeitet werden. Die Verarbeitung findet dadurch statt, daß der Vektor der zwei zeitdiskreten Abtastwerte aus unterschiedlichen Vierteln des Fensters mit einer Drehmatrix beaufschlagt werden, wobei die Drehmatrix durch eine Mehrzahl von sogenannten Lifting-Matrizen darstellbar ist. Wie es bekannt ist, zeichnen sich Lifting-Matrizen dadurch aus, daß sie lediglich ein Element haben, das ungleich "0" oder "1" ist, d. h. das nicht-

- ganzzahlig ist. Durch sequentielles Ausführen einer Multiplikation einer Lifting-Matrix mit dem Vektor von zeitdiskreten Abtastwerten und anschließendes Runden der Komponente des Vektors, die nicht-ganzzahlig ist, werden Gleitkommazahlen unmittelbar nach ihrer Entstehung gleich gerundet.
- 5 Es sei darauf hingewiesen, daß aufgrund der beschriebenen Eigenschaft der Lifting-Matrizen immer nur eine Komponente des Ergebnisvektors der Multiplikation gerundet werden muß.
- 10 Vorzugsweise ist die Drehmatrix eine Givens-Drehmatrix, die, wie es bekannt ist, durch drei Lifting-Matrizen darstellbar ist. Der Drehwinkel der Givens-Drehmatrix hängt von der Fensterfunktion ab. Es sei darauf hingewiesen, daß für das erfindungsgemäße Verfahren sämtliche Fensterfunktionen
- 15 zulässig sind, so lange diese Fensterfunktionen die beschriebene Bedingung erfüllen, daß das Quadrat eines Fenstergewichts und das Quadrat eines um N Fenstergewichte zeitlich entfernten Fenstergewichts zusammen den Wert 1 ergeben. Es sei darauf hingewiesen, daß diese Bedingung auch
- 20 durch zwei aufeinanderfolgende Fenster unterschiedlicher Form, wie z. B. ein Sinus-Fenster und ein Kaiser-Bessel-Fenster, erfüllt sein kann.

Bei einem bevorzugten Ausführungsbeispiel der vorliegenden

25 Erfindung wird die MDCT-Verarbeitung mit 50%iger Überlappung durch Lifting-Matrizen und Rundungen und eine anschließende diskrete Cosinustransformation (DCT) mit nicht-symmetrischen Basisfunktionen, d. h. durch eine DCT vom Typ IV, ersetzt.

30

Um nicht nur eine ganzzahlige Fensterung sondern auch eine ganzzahlige diskrete Cosinustransformation zu erreichen, wird es ferner bevorzugt, auch die DCT-Transformation durch Givens-Rotationen und insbesondere durch Verarbeitung mit

35 Lifting-Matrizen und Rundung nach jeder Lifting-Matrix-Multiplikation zu ersetzen.

Ein Vorteil der vorliegenden Erfindung besteht darin, daß nunmehr entweder beim Fenstern oder beim kompletten Transformieren die Ausgangswerte, also die gefensterten Abtastwerte oder die Spektralwerte ganzzahlig bleiben. Dennoch ist der gesamte Prozeß umkehrbar, indem einfach die inversen Rotationen in umgekehrter Reihenfolge bezüglich der Verarbeitung der Lifting-Matrizen und dieselbe Rundungsfunktion eingesetzt werden. Das erfindungsgemäße Konzept eignet sich somit als ganzzahlige Approximierung der MDCT mit perfekter Rekonstruierbarkeit und stellt somit eine ganzzahlige modifizierte diskrete Cosinustransformation (INT MDCT) dar.

Das erfindungsgemäße Konzept hat nach wie vor die günstigen Eigenschaften der MDCT, d. h. eine überlappende Struktur, die eine bessere Frequenzselektivität als nicht-überlappende Blocktransformationen liefert, und eine kritische Abtastung, so daß die Gesamtanzahl von Spektralwerten, die ein Audiosignal darstellen, nicht die Anzahl von Eingang-Abtastwerten überschreitet. So werden zwar aufgrund der Rundungen in den Rotationsschritten Nichtlinearitäten eingeführt. Die Rundungen führen jedoch gleichzeitig dazu, daß der Zahlenbereich der ganzzahligen Spektralwerte den Zahlenbereich der Eingangswerte nicht wesentlich überschreitet. Obgleich aufgrund der überlappenden Struktur eine Energieerhaltung auf einer Block-um-Block-Basis, wie sie durch das Parsevalsche Theorem gegeben ist, nicht vorliegt, zeichnet sich die erfindungsgemäße ganzzahlige MDCT dadurch aus, daß die mittlere Energie pro Block beibehalten wird, da vorzugsweise lediglich gerundete Givens-Rotationen eingesetzt werden, die im allgemeinen energieerhaltend sind.

Ein Vorteil der vorliegenden Erfindung besteht ferner darin, daß aufgrund der Tatsache, daß ganzzahlige Ausgangswerte vorliegen, auf eine anschließende Quantisierung verzichtet werden kann, derart, daß die Ausgangswerte der ganzzahligen MDCT unmittelbar Entropie-codiert werden können, um einen verlustlosen Datenkomprimierer zu erhalten.

Bevorzugte Ausführungsbeispiele der vorliegenden Erfindung werden nachfolgend Bezug nehmend auf die beiliegenden Zeichnungen detailliert erläutert. Es zeigen:

5

Fig. 1 ein Blockschaltbild einer erfindungsgemäßen Vorrichtung zum Verarbeiten von zeitdiskreten Audio-Abtastwerten, um ganzzahlige Werte zu erhalten;

10

Fig. 2 eine schematische Darstellung der Zerlegung einer MDCT und einer inversen MDCT in Givens-Rotationen und zwei DCT-IV-Operationen gemäß einem bevorzugten Ausführungsbeispiel der vorliegenden Erfindung;

15

Fig. 3 eine Darstellung zur Veranschaulichung der Zerlegung der MDCT mit 50-Prozent-Überlappung in Rotationen und DCT-IV-Operationen;

20

Fig. 4a ein schematisches Blockschaltbild eines bekannten Codierers mit MDCT und 50 Prozent Überlappung; und

25

Fig. 4b ein Blockschaltbild eines bekannten Decodierers zum Decodieren der durch Fig. 4a erzeugten Werte.

30

Fig. 1 zeigt ein Übersichtsdiagramm für die erfindungsgemäße Vorrichtung bzw. das erfindungsgemäße Verfahren zum Verarbeiten von zeitdiskreten Abtastwerten, die ein Audiosignal darstellen, um ganzzahlige Werte zu erhalten. Die zeitdiskreten Abtastwerte werden durch die in Fig. 1 gezeigte Vorrichtung gefenstert und optional in eine spektrale Darstellung umgesetzt. Die zeitdiskreten Abtastwerte, die an einem Eingang 10 in die Vorrichtung eingespeist werden, werden mit einem Fenster w mit einer Länge, die $2N$ zeitdiskreten Abtastwerten entspricht, gefenstert, um an einem Ausgang 12 ganzzahlige gefensterte Abtastwerte zu erreichen, welche dazu geeignet sind, um mittels einer Trans-

35

formation und insbesondere der Einrichtung 14 zum Ausführen einer ganzzahligen DCT in eine spektrale Darstellung umgesetzt zu werden. Die ganzzahlige DCT ist ausgebildet, um aus N Eingangswerten N Ausgangswerte zu erzeugen, was im
5 Gegensatz zu der MDCT-Funktion 408 von Fig. 4a steht, die aus 2N gefensterten Abtastwerten aufgrund der MDCT-Gleichung lediglich N Spektralwerte erzeugt.

10 Zum Fenstern der zeitdiskreten Abtastwerte werden zunächst in einer Einrichtung 16 zwei zeitdiskrete Abtastwerte ausgewählt, die zusammen einen Vektor von zeitdiskreten Abtastwerten darstellen. Ein zeitdiskreter Abtastwert, der durch die Einrichtung 16 ausgewählt wird, liegt im ersten Viertel des Fensters. Der andere zeitdiskrete Abtastwert
15 liegt in dem zweiten Viertel des Fensters, wie es anhand von Fig. 3 noch detaillierter ausgeführt wird. Der durch die Einrichtung 16 erzeugte Vektor wird nunmehr mit einer Drehmatrix der Dimension 2×2 beaufschlagt, wobei diese Operation nicht unmittelbar durchgeführt wird, sondern mit-
20 tels mehrerer sogenannten Lifting-Matrizen.

Eine Lifting-Matrix hat die Eigenschaft, daß sie nur ein Element aufweist, das von dem Fenster w abhängt und ungleich "1" oder "0" ist.

25 Die Faktorisierung von Wavelet-Transformationen in Lifting-Schritte ist in der Fachveröffentlichung "Factoring Wavelet Transforms Into Lifting Steps", Ingrid Daubechies und Wim Sweldens, Preprint, Bell Laboratories, Lucent Technologies, 1996, dargestellt. Allgemein ist ein Lifting-Schema eine
30 einfache Beziehung zwischen perfekt rekonstruierenden Filterpaaren, die dasselbe Tiefpaß- oder Hochpaßfilter haben. Jedes Paar komplementärer Filter kann in Lifting-Schritte faktorisiert werden. Insbesondere gilt dies für Givens-
35 Rotationen. Es sei der Fall betrachtet, bei dem die Polyphasenmatrix eine Givens-Rotation ist. Es gilt dann:

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} = \begin{pmatrix} 1 & \frac{\cos \alpha - 1}{\sin \alpha} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \sin \alpha & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{\cos \alpha - 1}{\sin \alpha} \\ 0 & 1 \end{pmatrix}$$

Jede der drei rechts des Gleichheitszeichens stehenden Lifting-Matrizen hat als Hauptdiagonalelemente den Wert "1".

- 5 Ferner ist in jeder Lifting-Matrix ein Nebendiagonalelement gleich 0, und ein Nebendiagonalelement vom Drehwinkel α abhängig.

Der Vektor wird nunmehr mit der dritten Lifting-Matrix, d.
 10 h. der Liftingmatrix ganz rechts in obiger Gleichung, multipliziert, um einen ersten Ergebnisvektor zu erhalten. Dies ist in Fig. 1 durch eine Einrichtung 18 dargestellt. Erfindungsgemäß wird nunmehr der erste Ergebnisvektor mit einer beliebigen Rundungsfunktion, die die Menge der reel-
 15 len Zahlen in die Menge der ganzen Zahlen abbildet, gerundet, wie es in Fig. 1 durch eine Einrichtung 20 dargestellt ist. Am Ausgang der Einrichtung 20 wird ein gerundeter erster Ergebnisvektor erhalten. Der gerundete erste Ergebnisvektor wird nunmehr in eine Einrichtung 22 zum Multiplizieren desselben mit der mittleren, d. h. zweiten, Lifting-
 20 Matrix eingespeist, um einen zweiten Ergebnisvektor zu erhalten, der in einer Einrichtung 24 wiederum gerundet wird, um einen gerundeten zweiten Ergebnisvektor zu erhalten. Der gerundete zweite Ergebnisvektor wird nunmehr in eine Ein-
 25 richtung 26 eingespeist, und zwar zum Multiplizieren desselben mit der links in der obigen Gleichung aufgeführten, d. h. ersten, Liftingmatrix, um einen dritten Ergebnisvektor zu erhalten, der schließlich noch mittels einer Ein-
 richtung 28 gerundet wird, um schließlich an dem Ausgang 12
 30 ganzzahlige gefensterte Abtastwerte zu erhalten, die nun, wenn eine spektrale Darstellung derselben gewünscht wird, durch die Einrichtung 14 verarbeitet werden müssen, um an einem Spektralausgang 30 ganzzahlige Spektralwerte zu erhalten.

35

Vorzugsweise ist die Einrichtung 14 als Ganzzahl-DCT oder Integer-DCT ausgeführt.

Die diskrete Cosinus-Transformation gemäß Typ 4 (DCT-IV) mit einer Länge N ist durch folgende Gleichung gegeben:

$$X_t(m) = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} x(k) \cos\left(\frac{\pi}{4N}(2k+1)(2m+1)\right)$$

5

Die Koeffizienten der DCT-IV bilden eine orthonormale N x N Matrix. Jede orthogonale N x N-Matrix kann in N (N-1)/2 Givens-Rotationen zerlegt werden, wie es in der Fachveröffentlichung P. P. Vaidyanathan, "Multirate Systems And Filter Banks", Prentice Hall, Englewood Cliffs, 1993, ausgeführt ist. Es sei darauf hingewiesen, daß auch weitere Zerlegungen existieren.

Bezüglich der Klassifikationen der verschiedenen DCT-
15 Algorithmen sei auf H. S. Malvar, "Signal Processing With Lapped Transforms", Artech House, 1992, verwiesen. Allgemein unterscheiden sich die DCT-Algorithmen durch die Art ihrer Basisfunktionen. Während die DCT-IV, die bei der vorliegenden Erfindung bevorzugt wird, nicht-symmetrische Basisfunktionen umfaßt, d. h. eine Cosinus-Viertelwelle, eine Cosinus-3/4-Welle, eine Cosinus-5/4-Welle, eine Cosinus-7/4-Welle, etc., hat die diskrete Cosinustransformation z. B. vom Typ II (DCT-II), achsensymmetrische und punktsymmetrische Basisfunktionen. Die 0-te Basisfunktion hat einen
20 Gleichanteil, die erste Basisfunktion ist eine halbe Cosinuswelle, die zweite Basisfunktion ist eine ganze Cosinuswelle, usw. Aufgrund der Tatsache, daß die DCT-II den Gleichanteil besonders berücksichtigt, wird sie bei der Videocodierung verwendet, nicht aber bei der Audiocodierung, da bei der Audiocodierung im Gegensatz zur Videocodierung der Gleichanteil nicht von Relevanz ist.
30

35

Im nachfolgend wird darauf eingegangen, wie der Drehwinkel α der Givens-Rotation von der Fensterfunktion abhängt.

Eine MDCT mit einer Fensterlänge von 2N kann in eine diskrete Cosinustransformation vom Typ IV mit einer Länge N reduziert werden. Dies wird dadurch erreicht, daß die TDAC-

Operation explizit im Zeitbereich durchgeführt wird, und daß dann die DCT-IV angewandt wird. Bei einer 50%igen Überlappung überlappt die linke Hälfte des Fensters für einen Block t mit der rechten Hälfte des vorausgehenden Blocks, d. h. des Blocks t-1. Der überlappende Teil zwei aufeinanderfolgender Blöcke t-1 und t wird im Zeitbereich, d. h. vor der Transformation, folgendermaßen vorverarbeitet, d. h. zwischen dem Eingang 10 und dem Ausgang 12 von Fig. 1 verarbeitet:

10

$$\begin{pmatrix} \tilde{x}_t(k) \\ \tilde{x}_{t-1}(N-1-k) \end{pmatrix} = \begin{pmatrix} w(\frac{N}{2} + k) & -w(\frac{N}{2} - 1 - k) \\ w(\frac{N}{2} - 1 - k) & w(\frac{N}{2} + k) \end{pmatrix} \begin{pmatrix} x_t(\frac{N}{2} + k) \\ x_t(\frac{N}{2} - 1 - k) \end{pmatrix}$$

15 Die mit der Tilde bezeichneten Werte sind die Werte am Ausgang 12 von Fig. 1, während die ohne Tilde in der obigen Gleichung bezeichnete x Werte die Werte am Eingang 10 bzw. hinter der Einrichtung 16 zum Auswählen sind. Der Laufindex k läuft von 0 bis N/2-1, während w die Fensterfunktion dar-

20 stellt.

Aus der TDAC-Bedingung für die Fensterfunktion w gilt folgender Zusammenhang:

25

$$w(\frac{N}{2} + k)^2 + w(\frac{N}{2} - 1 - k)^2 = 1$$

Für bestimmte Winkel α_k , $k = 0, \dots, N/2-1$ kann diese Vor-

30 verarbeitung im Zeitbereich als Givens-Rotation geschrieben werden, wie es ausgeführt worden ist.

Der Winkel α der Givens-Rotation hängt folgendermaßen von der Fensterfunktion w ab:

35

$$\alpha = \arctan[w(N/2-1-k) / w(N/2 + k)]$$

Es sei darauf hingewiesen, daß beliebige Fensterfunktionen w eingesetzt werden können, solange sie diese TDAC-Bedingung erfüllen.

- 5 Im nachfolgenden wird anhand von Fig. 2 ein kaskadierter Codierer und Decodierer beschrieben. Die zeitdiskreten Abtastwerte $x(0)$ bis $x(2N-1)$, die durch ein Fenster gemeinsam "gefenstert" werden, werden zunächst derart durch die Einrichtung 16 von Fig. 1 ausgewählt, daß der Abtastwert $x(0)$
- 10 und der Abtastwert $x(N-1)$, d. h. ein Abtastwert aus dem ersten Viertel des Fensters und ein Abtastwert aus dem zweiten Viertel des Fensters, ausgewählt werden, um den Vektor am Ausgang der Einrichtung 16 zu bilden. Die sich überkreuzenden Pfeile stellen schematisch die Lifting-
- 15 Multiplikationen und anschließenden Rundungen der Einrichtungen 18, 20 bzw. 22, 24 bzw. 26, 28 dar, um am Eingang der DCT-IV-Blöcke die ganzzahligen gefensterten Abtastwerte zu erhalten.
- 20 Wenn der erste Vektor wie oben beschrieben verarbeitet ist, wird ferner ein zweiter Vektor aus den Abtastwerten $x(N/2-1)$ und $x(N/2)$, d. h. wieder ein Abtastwert aus dem ersten Viertel des Fenster und ein Abtastwert aus dem zweiten Viertel des Fensters, ausgewählt und wiederum durch den in
- 25 Fig. 1 beschriebenen Algorithmus verarbeitet. Analog dazu werden sämtliche anderen Abtastwertpaare aus dem ersten und zweiten Viertel des Fensters bearbeitet. Die selbe Verarbeitung wird für das dritte und vierte Viertel des ersten Fensters durchgeführt. Nunmehr liegen am Ausgang 12 $2N$ gefensterte ganzzahlige Abtastwerte vor, die nunmehr so, wie
- 30 es in Fig. 2 dargestellt ist, in eine DCT-IV-Transformation eingespeist werden. Insbesondere werden die ganzzahligen gefensterten Abtastwerte des zweiten und dritten Viertels in eine DCT eingespeist. Die gefensterten ganzzahligen Ab-
- 35 tastwerte des ersten Viertels des Fensters werden in eine vorausgehende DCT-IV zusammen mit den gefensterten ganzzahligen Abtastwerten des vierten Viertels des vorausgehenden Fensters verarbeitet. Analog dazu wird das vierte Viertel

der gefensterten ganzzahligen Abtastwerte in Fig. 2 mit dem ersten Viertel des nächsten Fensters zusammen in eine DCT-IV-Transformation eingespeist. Die mittlere in Fig. 2 gezeigte ganzzahlige DCT-IV-Transformation 32 liefert nunmehr
5 N ganzzahlige Spektralwerte $y(0)$ bis $y(N-1)$. Diese ganzzahligen Spektralwerte können nunmehr beispielsweise einfach Entropie-codiert werden, ohne daß eine dazwischenliegende Quantisierung erforderlich ist, da die erfindungsgemäße Fensterung und Transformation ganzzahlige Ausgangswerte
10 liefert.

In der rechten Hälfte von Fig. 2 ist ein Decodierer dargestellt. Der Decodierer bestehend aus Rücktransformation und "inverser Fensterung" arbeitet invers zum Codierer. Es ist
15 bekannt, daß zur Rücktransformation einer DCT-IV eine inverse DCT-IV verwendet werden kann, wie es in Fig. 2 dargestellt ist. Die Ausgangswerte der Decodierer-DCT-IV 34 werden nunmehr, wie es in Fig. 2 dargestellt ist, mit den entsprechenden Werten der vorausgehenden Transformation bzw.
20 der nachfolgenden Transformation gemäß der vorliegenden Erfindung invers verarbeitet, um aus den ganzzahligen gefensterten Abtastwerten am Ausgang der Einrichtung 34 bzw. der vorausgehenden und nachfolgenden Transformation wieder zeitdiskrete Audio-Abtastwerte $x(0)$ bis $x(2N-1)$ zu erzeugen.
25

Die ausgangsseitige Operation geschieht erfindungsgemäß durch eine inverse Givens-Rotation, d. h. derart, daß die Blöcke 26, 28 bzw. 22, 24 bzw. 18, 20 in der entgegengesetzten Richtung durchlaufen werden. Dies sei anhand der
30 zweiten Lifting-Matrix von Gleichung 1 näher dargestellt. Wenn (im Codierer) der zweite Ergebnisvektor durch Multiplikation des gerundeten ersten Ergebnisvektors mit der zweiten Liftingmatrix (Einrichtung 22) gebildet wird, so
35 ergibt sich folgender Ausdruck:

$$(x, y) \mapsto (x, y + x \sin \alpha)$$

Die Werte x , y auf der rechten Seite von Gleichung 6 sind
 Ganzzahlen. Dies trifft jedoch für den Wert $x \sin \alpha$ nicht
 zu. Hier muß die Rundungsfunktion r eingeführt werden, wie
 5 es in der nachfolgenden Gleichung

$$(x, y) \mapsto (x, y + r(x \sin \alpha))$$

10

dargestellt ist. Diese Operation führt die Einrichtung 24
 aus.

Die inverse Abbildung (im Decodierer) ist folgendermaßen
 15 definiert:

$$(x', y') \mapsto (x', y' - r(x' \sin \alpha))$$

20

Aufgrund dem Minuszeichens vor der Rundungsoperation wird
 ersichtlich, daß die ganzzahlige Approximierung des Lifting-Schritts
 umgekehrt werden kann, ohne daß ein Fehler eingeführt wird. Die
 Anwendung dieser Approximation auf je-
 25 den der drei Lifting-Schritte führt zu einer ganzzahligen
 Approximation der Givens-Rotation. Die gerundete Rotation
 (im Codierer) kann umgekehrt werden (im Decodierer), ohne
 daß ein Fehler eingeführt wird, und zwar indem die inversen
 gerundeten Lifting-Schritte in umgekehrter Reihenfolge
 30 durchlaufen werden, d. h. wenn beim Decodieren der Algo-
 rithmus von Fig. 1 von unten nach oben durchgeführt wird.

Wenn die Rundungsfunktion r punktsymmetrisch ist, ist die
 inverse gerundete Rotation identisch zu der gerundeten Ro-
 35 tation mit dem Winkel $-\alpha$ und lautet folgendermaßen:

$$\begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$$

Die Lifting-Matrizen für den Decodierer, d. h. für die inverse Givens-Rotation, ergibt sich in diesem Fall unmittelbar aus Gleichung (1), indem lediglich der Ausdruck "sin α " durch den Ausdruck "-sin α " ersetzt wird.

Im nachfolgenden wird anhand von Fig. 3 noch einmal die Zerlegung einer üblichen MDCT mit überlappenden Fenstern 40 bis 46 dargelegt. Die Fenster 40 bis 46 überlappen jeweils zu 50%. Pro Fenster werden zunächst Givens-Rotationen innerhalb des ersten und zweiten Viertels eines Fensters bzw. innerhalb des dritten und vierten Viertels eines Fensters ausgeführt, wie es durch die Pfeile 48 schematisch darstellt ist. Dann werden die rotierten Werte, d. h. die gefensterten ganzzahligen Abtastwerte derart in eine N-zu-N-DCT eingespeist, daß immer das zweite und dritte Viertel eines Fensters bzw. das vierte und erste Viertel eines darauffolgenden Fensters gemeinsam mittels eines DCT-IV-Algorithmus in eine spektrale Darstellung umgesetzt wird.

Erfindungsgemäß werden daher die üblichen Givens-Rotation in Lifting-Matrizen zerlegt, die sequentiell ausgeführt werden, wobei nach jeder Lifting-Matrix-Multiplikation ein Rundungsschritt eingeführt wird, derart, daß die Gleitkommazahlen unmittelbar nach ihrer Entstehung gerundet werden, derart, daß vor jeder Multiplikation eines Ergebnisvektors mit einer Lifting-Matrix der Ergebnisvektor lediglich Ganzzahlen hat.

30

Die Ausgangswerte bleiben also immer ganzzahlig, wobei es bevorzugt wird, auch ganzzahlige Eingangswerte zu verwenden. Dies stellt keine Einschränkung dar, da jegliche beispielsweise PCM-Abtastwerte, wie sie auf einer CD abgespeichert sind, ganzzahlige Zahlenwerte sind, deren Wertebereich je nach Bitbreite variiert, d. h. abhängig davon, ob die zeitdiskreten digitalen Eingangswerte 16-Bit-Werte oder 24-Bit-Werte sind. Dennoch ist, wie es ausgeführt worden

35

ist, der gesamte Prozeß invertierbar, indem die inversen Rotationen in umgekehrter Reihenfolge ausgeführt werden. Erfindungsgemäß existiert somit eine ganzzahlige Approximation der MDCT mit perfekter Rekonstruktion, also eine verlustlose Transformation.

Die erfindungsgemäße Transformation liefert ganzzahlige Ausgangswerte statt Gleitkommawerte. Sie liefert eine perfekte Rekonstruktion, so daß kein Fehler eingeführt wird, wenn eine Vorwärts- und dann eine Rückwärtstransformation ausgeführt werden. Die Transformation ist gemäß einem bevorzugten Ausführungsbeispiel der vorliegenden Erfindung ein Ersatz für die modifizierte diskrete Cosinustransformation. Auch andere Transformationsverfahren können jedoch ganzzahlig ausgeführt werden, so lange eine Zerlegung in Rotationen und eine Zerlegung der Rotationen in Lifting-Schritte möglich ist.

Die ganzzahlige MDCT gemäß der vorliegenden Erfindung hat die meisten günstigen Eigenschaften der MDCT. Sie hat eine überlappende Struktur, wodurch eine bessere Frequenzselektivität als bei nicht-überlappenden Blocktransformationen erhalten wird. Aufgrund der TDAC-Funktion, die bereits beim Fenstern vor der Transformation berücksichtigt wird, wird eine kritische Abtastung beibehalten, so daß die Gesamtanzahl von Spektralwerten, die ein Audiosignal darstellen, gleich der Gesamtanzahl von Eingangs-Abtastwerten ist.

Verglichen mit einer normalen MDCT, die Gleitkomma-Abtastwerte liefert, zeigt sich bei der erfindungsgemäßen ganzzahligen Transformation, daß lediglich in dem Spektralbereich, in dem wenig Signalpegel ist, das Rauschen im Vergleich zur normalen MDCT erhöht ist, während sich diese Rauscherhöhung bei signifikanten Signalpegeln nicht bemerkbar macht. Dafür bietet sich die erfindungsgemäße ganzzahlige Verarbeitung für eine effiziente Hardware-Implementation an, da lediglich Multiplikationsschritte verwendet werden, die ohne weiteres in Verschieben-

Addieren-Schritte (Shift/Add-Schritte) zerlegt werden können, welche einfach und schnell hardwaremäßig implementiert werden können.

- 5 Die erfindungsgemäße ganzzahlige Transformation liefert eine gute spektrale Darstellung des Audiosignals und bleibt dennoch im Bereich der ganzen Zahlen. Wenn sie auf tonale Teile eines Audiosignals angewandt wird, resultiert dies in einer guten Energiekonzentrierung. Damit kann ein effizientes verlustloses Codierschema aufgebaut werden, indem einfach die in Fig. 1 dargestellte erfindungsgemäße Fensterung/Transformation mit einem Entropiecodierer kaskadiert wird. Insbesondere ein gestapeltes Codieren (Stacked Coding) unter Verwendung von Escape-Werten, wie es in MPEG AAC eingesetzt wird, ist für die vorliegende Erfindung günstig. Es wird bevorzugt, alle Werte um eine bestimmte Potenz von zwei herunterzuskalieren, bis sie in eine erwünschte Codetabelle passen, und dann die weggelassenen niederstwertigen Bits zusätzlich zu codieren. Im Vergleich zu der Alternative der Verwendung von größeren Codetabellen ist die beschriebene Alternative hinsichtlich des Speicher-
10 verbrauchs zum Speichern der Codetabellen günstiger. Ein nahezu verlustloser Codierer könnte auch dadurch erhalten werden, daß einfach bestimmte der niederstwertigen Bits
15 weggelassen werden.
20
25

- Insbesondere für tonale Signale ermöglicht eine Entropie-Codierung der ganzzahligen Spektralwerte einen hohen Codiergewinn. Für transiente Teile des Signals ist der Codiergewinn niedrig, und zwar aufgrund des flachen Spektrums transienter Signale, d. h. aufgrund einer geringen Anzahl von Spektralwerten, die gleich oder nahezu 0 sind. Wie es in J. Herre, J. D. Johnston: "Enhancing the Performance of Perceptual Audio Coders by Using Temporal Noise Shaping (TNS)" 101. AES Convention, Los Angeles, 1996, Preprint
30 4384, beschrieben ist, kann diese Flachheit jedoch verwendet werden, indem eine lineare Prädiktion im Frequenzbereich verwendet wird. Eine Alternative ist eine Prädiktion
35

mit offener Schleife. Eine andere Alternative ist der Prädiktor mit geschlossener Schleife. Die erste Alternative, d. h. der Prädiktor mit offener Schleife, wird TNS genannt. Die Quantisierung nach der Prädiktion führt zu einer Adaption des resultierenden Quantisierungsrauschens an die zeitliche Struktur des Audiosignals und verhindert daher Vorechos in psychoakustischen Audiocodierern. Für ein verlustloses Audiocodieren ist die zweite Alternative, d. h. mit einem Prädiktor mit geschlossener Schleife, geeigneter, da die Prädiktion mit geschlossener Schleife eine genaue Rekonstruktion des Eingangssignals erlaubt. Wenn diese Technik auf ein erfindungsgemäß erzeugtes Spektrum angewendet wird, muß ein Rundungsschritt nach jedem Schritt des Prädiktionsfilters durchgeführt werden, um im Bereich der Ganzzahlen zu bleiben. Durch Verwenden des inversen Filters und derselben Rundungsfunktion kann das ursprüngliche Spektrum genau wieder hergestellt werden.

Um die Redundanz zwischen zwei Kanälen zur Datenreduktion auszunutzen, kann auch eine Mitte-Seite-Codierung verlustlos eingesetzt werden, wenn eine gerundete Rotation mit einem Winkel $\pi/4$ verwendet wird. Im Vergleich zur Alternative des Berechnens der Summe und Differenz des linken und rechten Kanals eines Stereosignals hat die gerundete Rotation den Vorteil der Energieerhaltung. Die Verwendung sogenannter Joint-Stereo-Codiertechniken kann für jedes Band ein- oder ausgeschaltet werden, wie es auch im Standard MPEG AAC durchgeführt wird. Weitere Drehwinkel können ebenfalls berücksichtigt werden, um eine Redundanz zwischen zwei Kanälen flexibler reduzieren zu können.

Patentansprüche

1. Verfahren zum Verarbeiten von zeitdiskreten Abtastwerten, die ein Audiosignal darstellen, um ganzzahlige
5 Werte zu erhalten, mit folgenden Schritten:

Fenstern der zeitdiskreten Abtastwerte mit einem Fenster (w) mit einer Länge, die $2N$ zeitdiskreten Abtastwerten entspricht, um gefensterte zeitdiskrete Abtastwerte für eine Umsetzung der zeitdiskreten Abtastwerte in eine spektrale Darstellung mittels einer Transformation zu liefern, die aus N Eingangswerten N Ausgangswerte erzeugen kann, wobei das Fenster folgende Teilschritte aufweist:

15
Auswählen (16) eines zeitdiskreten Abtastwerts aus einem Viertel des Fensters und eines zeitdiskreten Abtastwerts aus einem anderen Viertel des Fensters, um einen Vektor von zeitdiskreten Abtastwerten zu erhalten;

20
Beaufschlagen des Vektors mit einer quadratischen Drehmatrix, deren Dimension mit der Dimension des Vektors übereinstimmt, wobei die Drehmatrix durch eine Mehrzahl von Lifting-Matrizen darstellbar ist, wobei eine Lifting-Matrix nur ein Element aufweist, das von dem Fenster (w) abhängt und ungleich 1 oder 0 ist, wobei der Teilschritt des Beaufschlagens folgende Unterschritte aufweist:

30
Multiplizieren (18) des Vektors mit einer Lifting-Matrix, um einen ersten Ergebnisvektor zu erhalten;

35
Runden (20) einer Komponente des ersten Ergebnisvektors mit einer Rundungsfunktion (r), die eine reelle Zahl auf eine ganze

Zahl abbildet, um einen gerundeten ersten Ergebnisvektor zu erhalten; und

5 sequentiell durchführen der Schritte des Multiplizierens (22) und Rundens (24) mit einer anderen Lifting-Matrix, bis alle Lifting-Matrizen abgearbeitet sind, um einen gedrehten Vektor zu erhalten, der einen ganzzahligen gefensterten Abtastwert aus dem
10 Viertel des Fensters und einen ganzzahligen gefensterten Abtastwert aus dem anderen Viertel des Fensters aufweist.

15 2. Verfahren nach Anspruch 1, bei dem die Lifting-Matrizen 2 x 2-Matrizen sind, und bei dem insgesamt drei Lifting-Matrizen pro Drehmatrix vorhanden sind.

20 3. Verfahren nach Anspruch 1 oder 2, bei dem der Schritt des Multiplizierens des Vektors oder eines gerundeten Ergebnisvektors durch Bilden von Partialprodukten und Aufsummieren der Partialprodukte durchgeführt wird, wobei der Schritt des Rundens mit nicht-ganzzahligen Partialprodukten durchgeführt wird, bevor aufsummiert wird.

25 4. Verfahren nach einem der vorhergehenden Ansprüche, bei dem das Fenster eine Anzahl $2N$ von Abtastwerten aufweist, die gleich einer Potenz der Basis 2 ist.

30 5. Verfahren nach einem der vorhergehenden Ansprüche, bei dem die Drehmatrix eine Givens-Drehmatrix ist.

35 6. Verfahren nach einem der vorhergehenden Schritte, das ferner folgenden Schritt aufweist:

Durchführen des Schritt des Fensters für alle zeitdiskreten Abtastwerte der verbleibenden Viertel des Fen-

sters, um $2N$ gefilterte ganzzahlige Abtastwerte zu erhalten; und

Umsetzen (14) von N gefensterten ganzzahligen Abtastwerten in eine spektrale Darstellung durch eine Ganzzahl-DCT für Werte mit den gefilterten ganzzahligen Abtastwerten des zweiten Viertels und des dritten Viertels des Fensters, um N ganzzahlige Spektralwerte zu erhalten.

7. Verfahren nach Anspruch 6, bei dem die Ganzzahl-DCT eine DCT ist, die nicht-symmetrische Basisfunktionen aufweist.

8. Verfahren nach Anspruch 7, bei dem die DCT eine Typ-IV-DCT ist.

9. Verfahren nach einem der Ansprüche 6 bis 8, bei dem die DCT in Givens-Rotationsmatrizen und diese wiederum in Lifting-Matrizen zerlegbar ist, wobei nach jeder Multiplikation mit einer Lifting-Matrix ein Rundungsschritt folgt.

10. Verfahren nach einem der vorhergehenden Ansprüche, bei dem die Givens-Drehmatrix folgende Gestalt hat:

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

bei dem die Lifting-Matrizen folgende Gestalt haben:

$$\begin{pmatrix} 1 & \frac{\cos \alpha - 1}{\sin \alpha} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \sin \alpha & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{\cos \alpha - 1}{\sin \alpha} \\ 0 & 1 \end{pmatrix}$$

bei dem der Winkel α folgendermaßen definiert ist:

$$\alpha = \arctan[w(N/2-1-k) / w(N/2 + k)]$$

5 wobei k ein Zeitindex der zeitdiskreten Abtastwerte ist und von 0 bis 2N-1 läuft und wobei w eine Fensterfunktion ist.

10 11. Verfahren nach einem der vorhergehenden Ansprüche, bei dem für die Fensterfunktion w folgende Bedingung erfüllt ist:

$$w\left(\frac{N}{2} + k\right)^2 + w\left(\frac{N}{2} - 1 - k\right)^2 = 1$$

15

12. Verfahren nach Anspruch 11, bei dem das Fenster ein Sinusfenster ist.

20 13. Verfahren nach einem der vorhergehenden Ansprüche, bei dem die zeitdiskreten Abtastwerte ganzzahlig sind.

14. Verfahren nach Anspruch 6, das ferner folgenden Schritt aufweist:

25

Entropie-Codieren der ganzzahligen Spektralwerte, um eine Entropie-codierte Version des Audiosignals zu erhalten.

30 15. Verfahren nach Anspruch 6, das ferner den Schritt des Quantisierens der ganzzahligen Spektralwerte unter Berücksichtigung der psychoakustischen Maskierungsschwelle aufweist, um quantisierte Spektralwerte zu erhalten, die so quantisiert sind, daß das Quantisierungsrauschen im wesentlichen maskiert ist.

35

16. Verfahren nach Anspruch 6, das ferner den Schritt des Filterns der ganzzahligen Spektralwerte über der Fre-

quenz mittels eines Prädiktors mit geschlossener Schleife und das Runden von Prädiktionsfehlerwerten aufweist.

- 5 17. Verfahren zum Invers-Verarbeiten von ganzzahligen Werten, die durch das Verfahren nach Anspruch 1 erzeugt worden sind, mit folgenden Schritten:

10 Beaufschlagen des gedrehten Vektors mit einer zu der Drehmatrix inversen Drehmatrix, wobei die inverse Drehmatrix durch eine Mehrzahl von Invers-Lifting-Matrizen darstellbar ist, wobei eine Invers-Lifting-Matrix nur ein Element aufweist, das von dem Fenster abhängt und ungleich 1 oder 0 ist, wobei der Schritt
15 des Beaufschlagens folgende Teilschritte aufweist:

Multiplizieren des gedrehten Vektors mit einer Invers-Lifting-Matrix, die zu der Lifting-Matrix invers ist, die beim Erzeugen der ganzzahligen
20 Werte zuletzt verwendet worden ist, um einen ersten Invers-Ergebnisvektor zu erhalten;

Runden einer Komponente des ersten Invers-Ergebnisvektors mit der Rundungsfunktion, um einen gerundeten ersten Invers-Ergebnisvektor zu erhalten; und
25

sequentielle Durchführen der Schritte des Multiplizierens und Rundens mit weiteren Lifting-Matrizen in einer Reihenfolge, die bezüglich der Reihenfolge beim Erzeugen der ganzzahligen Werte umgekehrt ist, um einen invers verarbeiteten Vektor zu erhalten, der einen ganzzahligen zeitdiskreten Abtastwert aus einem Viertel des Fensters
30 und einen ganzzahligen zeitdiskreten Abtastwert aus einem anderen Viertel des Fensters umfassen.
35

18. Verfahren nach Anspruch 17, bei dem die ganzzahligen Werte durch das Verfahren nach Anspruch 2 erzeugt sind und ganzzahlige Spektralwerte umfassen, und bei dem ferner vor dem Schritt des Beaufschlagens des gedrehten Vektors folgender Schritt ausgeführt wird:

Umsetzen der ganzzahligen Spektralwerte durch eine zu der Ganzzahl-DCT beim Erzeugen inverse Ganzzahl-DCT in eine zeitliche Darstellung, um den gedrehten Vektor zu erhalten.

19. Verfahren nach Anspruch 17 oder 18, bei dem die Rundungsfunktion punktsymmetrisch ist, und bei dem die Invers-Lifting-Matrizen identisch zu den Lifting-Matrizen sind, jedoch mit einem negativen Drehwinkel.

20. Verfahren nach Anspruch 18, bei dem die Ganzzahl-DCT eine von einer DCT vom Typ IV abgeleitete DCT ist.

21. Verfahren nach einem der Ansprüche 17 bis 20, bei dem die Invers-Lifting-Matrizen identisch zu den entsprechenden Lifting-Matrizen sind, jedoch abgesehen von Nebendiagonal-Elementen, die im Vergleich zu den Lifting-Matrizen bei den Invers-Lifting-Matrizen negativ sind.

22. Vorrichtung zum Verarbeiten von zeitdiskreten Abtastwerten, die ein Audiosignal darstellen, um ganzzahlige Werte zu erhalten, mit folgenden Merkmalen:

einer Einrichtung zum Fenstern der zeitdiskreten Abtastwerte mit einem Fenster (w) mit einer Länge, die $2N$ zeitdiskreten Abtastwerten entspricht, um gefensterter zeitdiskreter Abtastwerte für eine Umsetzung der zeitdiskreten Abtastwerte in eine spektrale Darstellung mittels einer Transformation zu liefern, die aus

N Eingangswerten N Ausgangswerte erzeugen kann, wobei die Einrichtung zum Fenstern folgende Teilmerkmale aufweist:

5 eine Einrichtung zum Auswählen (16) eines zeitdiskreten Abtastwerts aus einem Viertel des Fensters und eines zeitdiskreten Abtastwerts aus einem anderen Viertel des Fensters, um einen Vektor von zeitdiskreten Abtastwerten zu erhalten;

10 eine Einrichtung zum Beaufschlagen des Vektors mit einer quadratischen Drehmatrix, deren Dimension mit der Dimension des Vektors übereinstimmt, wobei die Drehmatrix durch eine Mehrzahl von Lifting-Matrizen darstellbar ist, wobei eine Lifting-Matrix nur ein Element aufweist, das von dem Fenster (w) abhängt und ungleich 1 oder 0 ist, wobei die Einrichtung zum Beaufschlagen folgende Untermerkmale aufweist:

20 eine Einrichtung zum Multiplizieren (18) des Vektors mit einer Lifting-Matrix, um einen ersten Ergebnisvektor zu erhalten;

25 eine Einrichtung zum Runden (20) einer Komponente des ersten Ergebnisvektors mit einer Rundungsfunktion (r), die eine reelle Zahl auf eine ganze Zahl abbildet, um einen gerundeten ersten Ergebnisvektor zu erhalten;

30 und

35 eine Einrichtung zum sequentiellen Durchführen der Schritte des Multiplizierens (22) und Rundens (24) mit einer anderen Lifting-Matrix, bis alle Lifting-Matrizen abgearbeitet sind, um einen gedrehten Vektor zu erhalten, der einen ganzzahligen gefensterten Abtastwert aus dem Viertel des Fensters und

einen ganzzahligen gefensterten Abtastwert aus dem anderen Viertel des Fensters aufweist.

- 5 23. Vorrichtung zum Invers-Verarbeiten von ganzzahligen Werten, die durch die Vorrichtung nach Anspruch 22 erzeugt worden sind, mit folgenden Merkmalen:

10 einer Einrichtung zum Beaufschlagen des gedrehten Vektors mit einer zu der Drehmatrix inversen Drehmatrix, wobei die inverse Drehmatrix durch eine Mehrzahl von Invers-Lifting-Matrizen darstellbar ist, wobei eine Invers-Lifting-Matrix nur ein Element aufweist, das von dem Fenster abhängt und ungleich 1 oder 0 ist, wobei die Einrichtung zum Beaufschlagen folgende Teilmerkmale aufweist:

20 eine Einrichtung zum Multiplizieren des gedrehten Vektors mit einer Invers-Lifting-Matrix, die zu der Lifting-Matrix invers ist, die beim Erzeugen der ganzzahligen Werte zuletzt verwendet worden ist, um einen ersten Invers-Ergebnisvektor zu erhalten;

25 eine Einrichtung zum Runden einer Komponente des ersten Invers-Ergebnisvektors mit der Rundungsfunktion, um einen gerundeten ersten Invers-Ergebnisvektor zu erhalten; und

30 eine Einrichtung zum sequentiellen Durchführen des Multiplizierens und Rundens mit weiteren Lifting-Matrizen in einer Reihenfolge, die bezüglich der Reihenfolge beim Erzeugen der ganzzahligen Werte umgekehrt ist, um einen invers verarbeiteten Vektor zu erhalten, der einen ganzzahligen zeitdiskreten Abtastwert aus einem Viertel des Fensters und einen ganzzahligen zeitdiskreten Ab-

35

tastwert aus einem anderen Viertel des Fensters
umfassen.

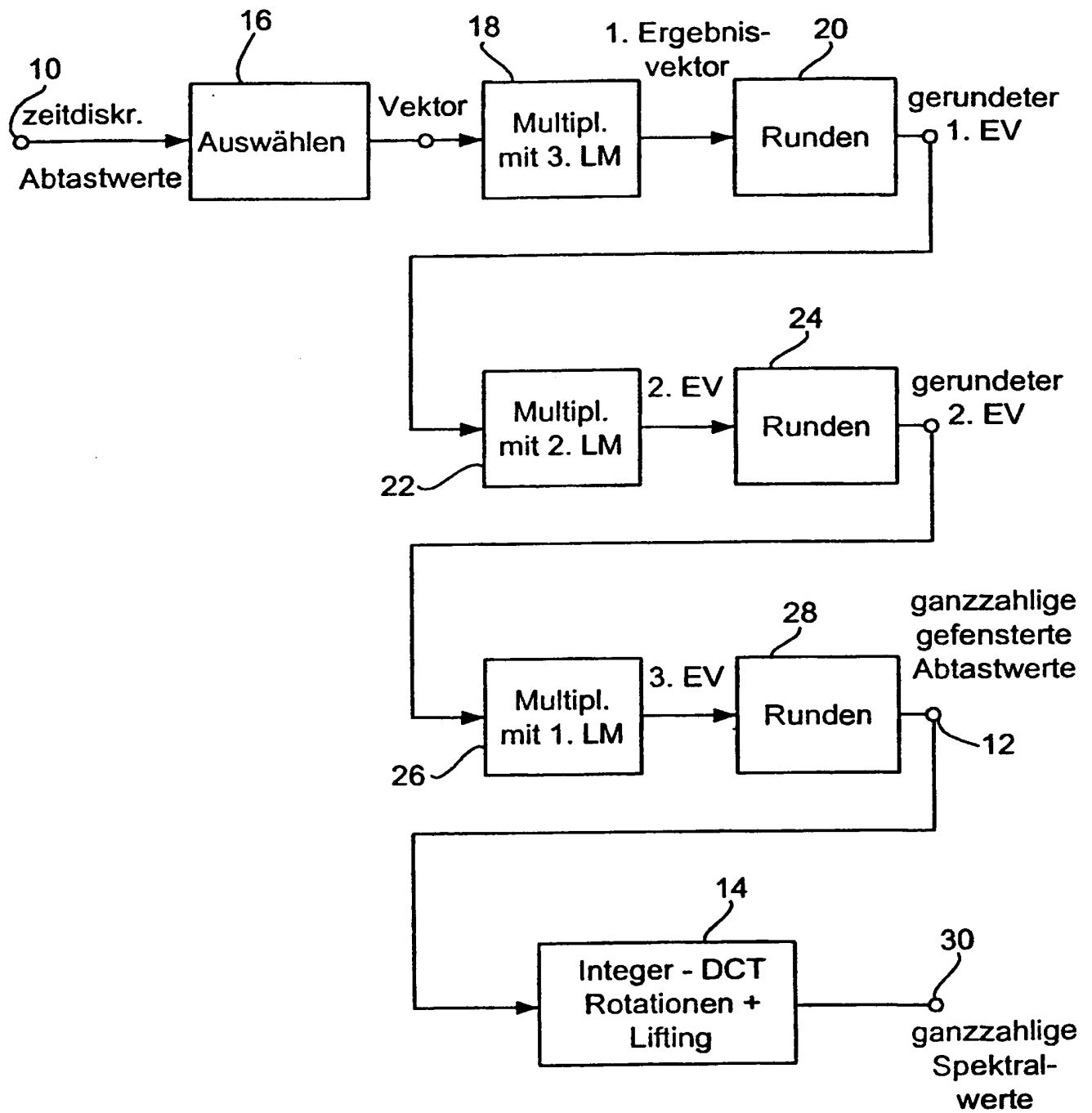


FIG. 1

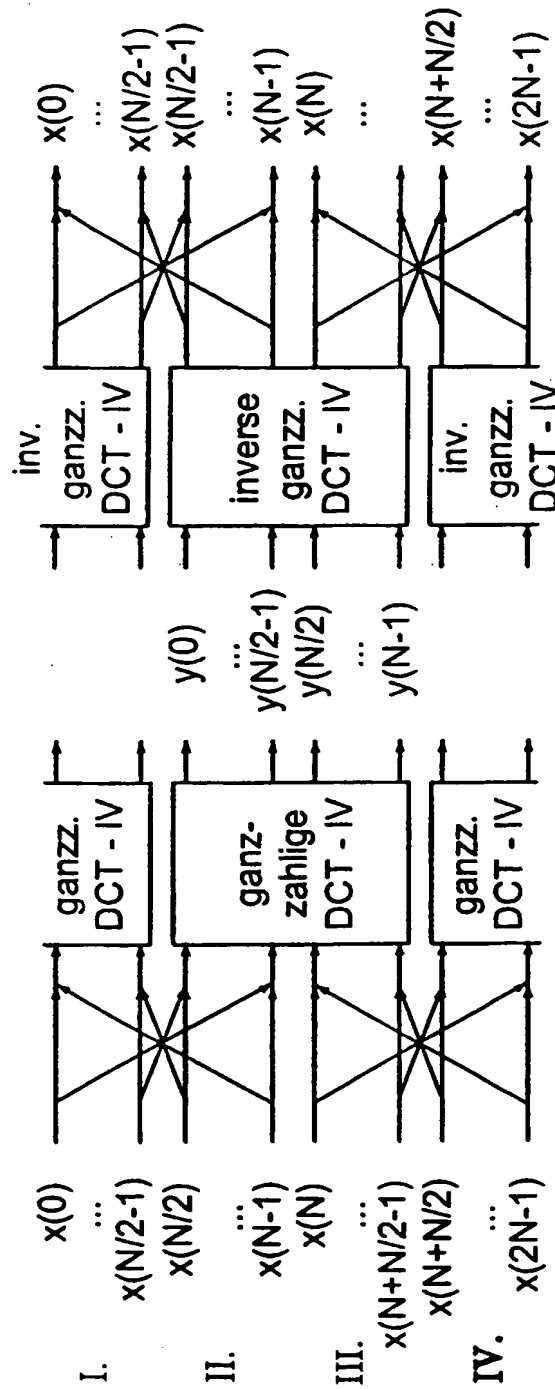


FIG. 2

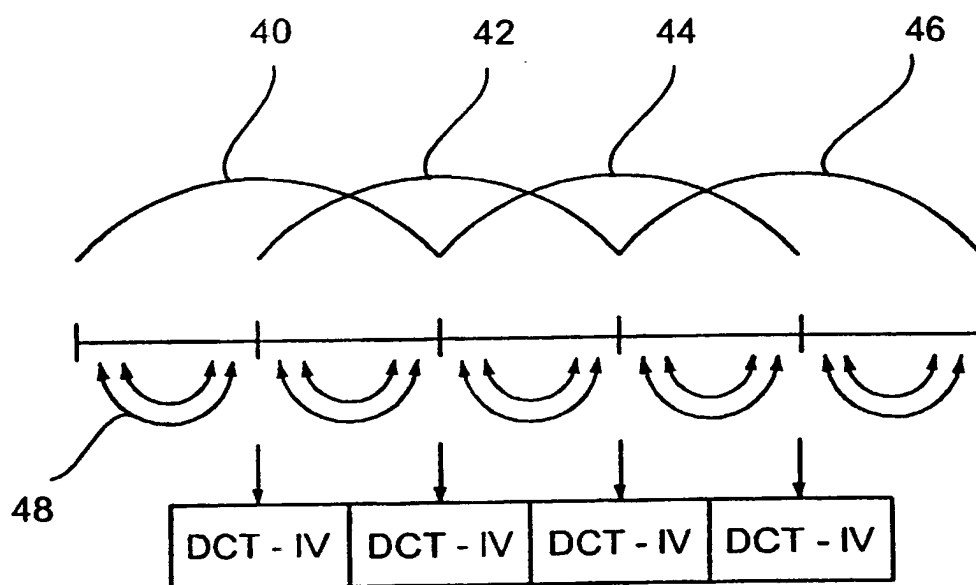


FIG. 3

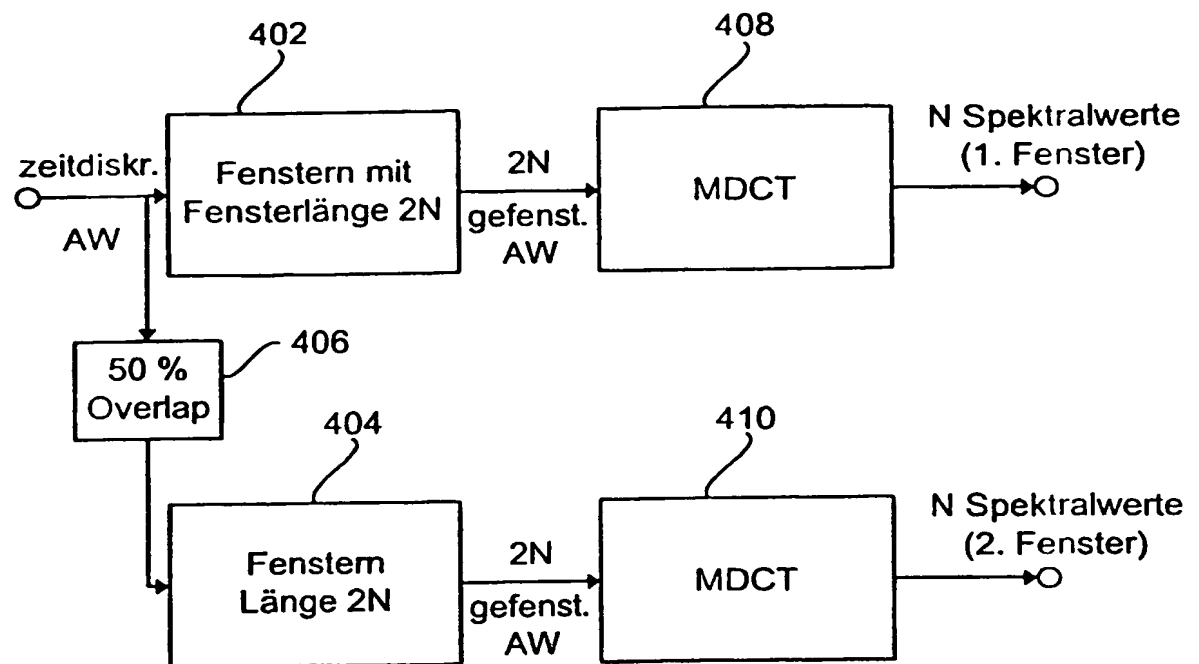


FIG. 4a

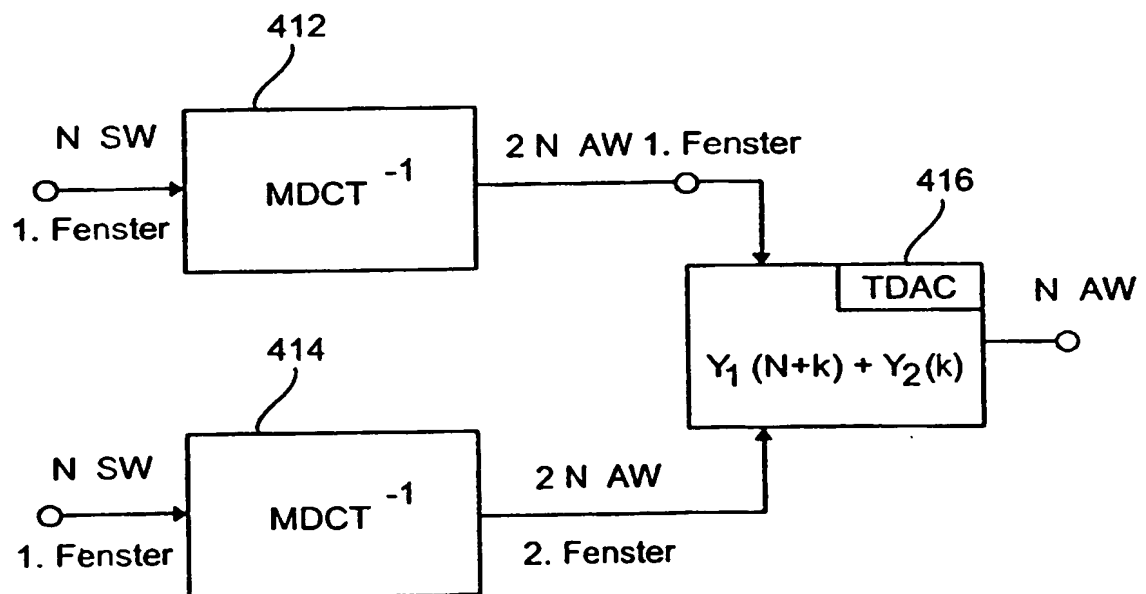


FIG. 4b

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 02/05865

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G10L19/02 G06F17/14

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G10L G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
INSPEC, COMPENDEX, WPI Data, EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ZENG Y ET AL: "Integer sinusoidal transforms based on lifting factorization" 2001 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. PROCEEDINGS (CAT. NO.01CH37221), 7 - 11 May 2001, pages 1181-1184 vol.2, XP002210315 SALT LAKE CITY, UT, USA, Piscataway, NJ, USA, IEEE, USA ISBN: 0-7803-7041-4 the whole document --- -/--	1-10, 13, 17-23

☒ Further documents are listed in the continuation of box C.

☐ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *G* document member of the same patent family

Date of the actual completion of the international search

26 August 2002

Date of mailing of the international search report

19/09/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.
Fax: (+31-70) 340-3016

Authorized officer

Quélavoine, R

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 02/05865

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>SOO-CHANG PEI ET AL: "The integer transforms analogous to discrete trigonometric transforms"</p> <p>IEEE TRANSACTIONS ON SIGNAL PROCESSING, DEC. 2000, IEEE, USA, vol. 48, no. 12, pages 3345-3364, XP002210316</p> <p>ISSN: 1053-587X</p> <p>abstract</p> <p>* Teil "V. Integer sine and cosine transforms" *</p> <p style="text-align: center;">---</p>	1-23
P,X	<p>GEIGER R., SPORER T., KOLLER J. AND BRANDENBURG K.: "Audio Coding based on Integer Transforms"</p> <p>111TH CONVENTION OF THE AUDIO ENGINEERING SOCIETY,</p> <p>no. 5471, 21 - 24 September 2001, XP008006797</p> <p>New York, NY, USA</p> <p>the whole document</p> <p style="text-align: center;">-----</p>	1-23

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP 02/05865

A. KLASSTIFIZIERUNG DES ANMELDUNGSGEGENSTANDES
IPK 7 G10L19/02 G06F17/14

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)
IPK 7 G10L G06F

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)
INSPEC, COMPENDEX, WPI Data, EPO-Internal

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	ZENG Y ET AL: "Integer sinusoidal transforms based on lifting factorization" 2001 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. PROCEEDINGS (CAT. NO.01CH37221), 7. - 11. Mai 2001, Seiten 1181-1184 vol.2, XP002210315 SALT LAKE CITY, UT, USA, Piscataway, NJ, USA, IEEE, USA ISBN: 0-7803-7041-4 das ganze Dokument --- -/--	1-10,13, 17-23

☒ Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen

☐ Siehe Anhang Patentfamilie

* Besondere Kategorien von angegebenen Veröffentlichungen :

A Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist

E älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

L Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)

O Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

P Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

T Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

X Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderscher Tätigkeit beruhend betrachtet werden

Y Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderscher Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

Z Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

26. August 2002

Absenddatum des internationalen Recherchenberichts

19/09/2002

Name und Postanschrift der internationalen Recherchenbehörde
Europäisches Patentamt, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+31-70) 340-3016

Bevollmächtigter Bediensteter

Quélavoine, R

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	SOO-CHANG PEI ET AL: "The integer transforms analogous to discrete trigonometric transforms" IEEE TRANSACTIONS ON SIGNAL PROCESSING, DEC. 2000, IEEE, USA, Bd. 48, Nr. 12, Seiten 3345-3364, XP002210316 ISSN: 1053-587X Zusammenfassung * Teil "V. Integer sine and cosine transforms" *	1-23
P,X	GEIGER R., SPORER T., KOLLER J. AND BRANDENBURG K.: "Audio Coding based on Integer Transforms" 111TH CONVENTION OF THE AUDIO ENGINEERING SOCIETY, Nr. 5471, 21. - 24. September 2001, XP008006797 New York, NY, USA das ganze Dokument	1-23

PRO PAGE RI ANK (USPTO)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☒ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)